

# ZJER

## ZIMBABWE JOURNAL OF EDUCATIONAL RESEARCH

### SPECIAL ISSUE

Volume 24 Number 2, July 2012

Technology Application in Primary Schools: Stakeholders views  
on the use of Calculators in Chinhoyi Urban.

*Emmanuel Chinamasa*

Teacher Evaluation by Pupils: Case of "O" Level Mathematics Student  
– Teachers in Bulawayo urban.

*Emmanuel Chinamasa, Morden Dzintotizeyi, Mathias Sithole*

Factors contributing to Teacher truancy in two Secondary Schools  
in Bulawayo. *Emmanuel Chinamasa, Ezekiel Sviye, Simbarashe Munikwa*

The Relevance of 'O' Level Mathematics in Nursing: A Survey  
of Practicing Nurses' Experiences in Zimbabwe.

*Matirwisa Kuneka, Emmanuel Chinamasa*

Secondary School Teachers' and Pupils' Views on the use of Mathematics  
Textbooks with Answers in Mazowe District.

*Lawrence Maregere, Emmanuel Chinamasa, Newton Hlenga*

Factors affecting Lecturer Research output in new Universities  
in Zimbabwe.

*Emmanuel Chinamasa*

Examinations Question Specialized Marking: A Quantitative Analysis  
of Inter-marker Reliability Mode at Chinhoyi University of Technology.

*Emmanuel Chinamasa, Cribert Munetsi*

Computation errors on measures of Central Tendency by Masters  
Students: Implications for Andragogy.

*Emmanuel Chinamasa, Cribert Munetsi*

Technology Utilisation: A survey of Computer Literacy levels among  
Health Personnel at Chinhoyi Provincial Hospital.

*Constance Madya, Emmanuel Chiramasa*

**The Zimbabwe Journal of Educational Research is published tri-annually by the University of Zimbabwe (UZ), Human Resources Research Centre (HRRC).**

**ISSN: 1013-3445**

**Editor-in-Chief: Professor Fred Zindi**

**Editorial Board**

**Prof. Levi M. Nyagura,  
University of Zimbabwe**

**Prof. V. Nyawaranda,  
University of Zimbabwe**

**Prof. Charles Nherera,  
Women's University in Africa**

**Prof. C. Mararike  
University of Zimbabwe**

**Editorial Advisory Board**

**Prof. Linda Chisholm  
Witwatersrand University**

**Prof. Danston S. J. Mkandawire  
University of Namibia**

**Prof John Schwillie  
Michigan State University**

***For further information contact us on:***

**Zimbabwe Journal of Educational Research  
HRRC, Faculty of Education  
University of Zimbabwe  
P. O. Box MP167  
Mount Pleasant  
HARARE  
Zimbabwe**

**E-mail: [hrrc@education.uz.ac.zw](mailto:hrrc@education.uz.ac.zw)**

**Tel: +263-04-303271 or 303211/9 Extn: 16002/3**

**ZJER**

**ZIMBABWE JOURNAL**

**OF**

**EDUCATIONAL RESEARCH**

## EXAMINATIONS QUESTION SPECIALIZED MARKING: A QUANTITATIVE ANALYSIS OF INTER-MARKER RELIABILITY MODE AT CHINHOYI UNIVERSITY OF TECHNOLOGY.

*Emmanuel Chinamasa, Chinhoyi University of Technology,  
Cribert Munetsi, University of Zimbabwe*

### Abstract

*The report presents findings from a quantitative quasi-experimental control and experiment group study. Its aim was to establish the effectiveness and advantages of examinations question specialized marking mode on inter-marker reliability. It was motivated by the current (2011) university problem of large volumes of marking against very few lecturers and the need to uphold a high level of quality marking. Data was gathered from two experimental marking sessions held in November 2010 and March 2011 at Chinhoyi University of Technology in Zimbabwe. Comparative analysis of deviations and time was done for the control and experiment group using paired t-tests, Spearman's rank order correlation and Chi-square for hypothesis testing. The study found that Question Specialised Marking (QSM) was effective. It reduced marker deviations, stress and marking time significantly. It has an inbuilt team member supervision mechanism, offers answer script security, reduces probability of marker malpractice and is fair to all students. The study recommends that, Universities with large marking loads and two or more lecturers per course can adopt Question Specialised Marking (QSM) for their next examination marking sessions to improve inter-marker reliability.*

**Key Words:** Examinations marking, Reliability, large classes, Quantitative analysis

### INTRODUCTION

Several, if not all universities observe that students can and often do ask for a script remark. It should be noted that the act of appealing against a failed examination result is an indirect acknowledgement of the fact that, marker reliability is low and that any studies like this one seeking ways of improving marker reliability in the face of an inverse relationship between the number of scripts, time and markers are welcome. At Zimbabwean Open University, appeals are paid for. A candidate is refunded if the appeal is successful. The appealing provision was put in place to improve transparency and fairness in the assessment of assignments and end of semester examinations of the essay type.

The assessment mark and grade that a candidate is awarded for any piece of work is crucial at University level. It is used to establish how much knowledge the learner has acquired from a series of learning activities, determining progress and motivating learners as proposed by Desforges (1990). To achieve the motivation, objective script marking should be reliable. Of great importance is the observation that, the teaching function of marked assignments rests in the way the marking is done and accepted by students as fair and rewarding what it should reward.

At school level Malmberg, Wanner, Sumra and little, (2000:128) found that, "the way teachers structure their teaching, provide feedback to pupils and awards grades are related to students' belief that they have effort, can put effort or are lucky." These individual deductions have specific bearing on students' habits, trust and respect for the education system and its qualifications.

It should be noted that, appeals are not the desired mode of operation. Universities would desire a zero tolerance based on accuracy, thoroughness, objectivity and promptness, which are the cardinal marker reliability indicators that this study aims to improve by applying question specialized marking. A high inter-marker reliability should be considered from a background in which essay and free response questions measuring higher order skills and application are preferred at tertiary level. A seeming contradiction arises from the fact that by their nature essay and free response questions have an inherent subjectivity in scoring. This is a lecturer function dependent on large classes and the time allocated for marking.

### **Contextual Framework**

The contextual situation of Zimbabwe's Universities requires an understanding of its historical perspectives. In Rhodesia education was the privilege of the elite in which a funneling academic education system for the black Africans was designed to cater for approximately 20% of the youths. That system left, "about 80% of potentially productive persons with no opportunity for training and acquiring employable skills." Chombo (2000:13).

After independence in 1980, the elected government declared education a human right and supported it by the education for all policy of 1987. The government observed the social demands for education from the electorate as spelt out by Mutumbuka (1981:7) who said, "it would be treachery indeed to betray the interests of the very people who have brought this government into power."

The implication of the education for all policy resulted in massive expansions at primary, secondary school levels and spilled into higher and tertiary education. Zimbabwe responded to this flood of student turn out at University by upgrading teachers' colleges into Universities and establishing other universities in other provinces. In this situation the country had to make do with limited lecturers who resisted the brain drain to green pastures syndrome that prevailed during the 2006 – 2009 galloping inflation.

The quality of students found in university today is an assorted bag in terms of their demographic distribution, intellectual capacity and orientation as a result of the prevailing political, social and economic factors. Specifically, the deteriorating social image of the teaching profession and public service in general compels some candidates "who would have been more profitably employed elsewhere" Moberly (1998:76), to be in the university. The closing down of the industrial sector also closed the route to apprenticeship, compelling other candidates who could have followed that path to divert to university. As if that was not enough, from the political spheres, affirmative action wheeled female students with weak passes into university.

As a result of this expansion, the country's budget could not manage. According to Williams in Zindi (1998:33), "In the 1988/89 financial year, 22% of the total government recurrent budget went to education." Economic realities required lecturers to generate income for their subsidiary functions. This did not go down well with lecturers who argued that, "their business is to teach and not fundraising" Chombo (2000:15). Reductions in government grants left lecturers with no better option but to introduce parallel and block-release programmes in which students are charged commercial rates for tuition fees. Four implications can be deduced from this move: first, two Zimbabwean candidates are buying the same education at different prices from the same institution. Second, those candidates with weak 'A' level passes but have the money are enrolled. The third is that working adults who have the money, need tertiary education and can spare the time gets into the University. Last but with a strong bearing on quality of instruction and examination script marking, is the observation that there is limited lecturer recruitment against massive student enrolment in order to maximize fees (income) and minimize salaries (expenditure).

From this position, one can justify the argument forwarded by Zindi (1998:33) who observed that, "...Because of large numbers, marking of assignments becomes an almost impossible task. Examination results consequently became unreliable as staff find it difficult to cope with large numbers." It is this maximum use of limited human resources situation which motivated this study to investigate the effects of applying question

specialized marking on inter-marker reliability as a survival strategy to ease the quantity burden and maintain a high quality of essay marking in universities.

In this study, Question Specialized Marking (QSM) refers to a process of marking assignments and examinations scripts in which, one marker marks only one question on each of the examination scripts. Bukenya (2006) called it belt marking. The scripts rotate from one marker to the next until all answers are marked. The last marker adds up the script totals. Inter-marker reliability refers to the consistency with which two or more independent markers award the same mark for the same essay.

### **Research Problem**

The study was prompted by the observation that, lecturers are required to mark piles and piles of assignments and examination scripts accurately, thoroughly, objectively and promptly within a short period of time to meet mark schedule submission dead lines. Practically and theoretically, the set up compromises the quality of marker reliability resulting in team leader or external examiner failing some candidates who would have been passed and passing others who would have been failed. In the past, cases of candidates who received failing results, paid for a re-mark and passed left stakeholders puzzled. The candidate's request for a script remarking is a reflection of the public's lack of confidence in the examination system, particularly the marking. It can be argued then that some candidates who lost a semester or year repeating a course were not supposed to. The scoring error had direct and indirect opportunity costs to the student, marker, institution and country whose productive human resource spent an extra year in college instead of being productive in the field. Such unreliable situations raise people's eyebrows over the credibility of the Zimbabwean Tertiary Education and credentials of the marker.

### **Research Questions**

The pathetic scenario described above, motivated three research questions:

1. How can question specialized marking be organized?
2. What are the effects of question specialized marking on inter-marker reliability?
3. What are participants' views of question specialized marking?

### **Hypothesis**

Three pairs of Hypothesis were generated to focus this study:

H<sub>0</sub>: Question specialized marking has no effect on inter-marker reliability.

H<sub>1</sub>: Question specialized marking improves inter-marker reliability.

$H_0$ : Question specialized marking has no effect on marking time.

$H_1$ : Question specialized marking reduces marking time, and

$H_0$ : Marking method is independent of marker

$H_1$ : Marking method depends on particular marker

### **Assumptions**

The study is driven by the following assumptions; Universities in Zimbabwe have large volumes of assignments and examinations scripts that are marked by at least two or more markers. Second, that lack of marker reliability is undesirable in any educational situation in which examinations play a pivotal role. Third is that, marker reliability is a variable normally distributed, whose effects in the examination system can be controlled by altering marking modalities. Last but important, examinations will serve their purpose if they are carefully drawn up, carefully marked and results presented in a meaningful way accepted by all stakeholders.

### **Significance of the Study**

This study is a quantitative action research with a high practical orientation, which derives its major significance from the following considerations:

1. It is probably one of the few, if not the first action research to provide possible solutions to the problem of low marker reliability at University level. This is an important contribution in Zimbabwe, where expansion is more inclined along the quantitative than qualitative dimension.
2. Study findings are practical, aiming to improve examinations marking practice. Its findings can be tried by any university during the next examination marking session. Other researchers can find the study rich in further research insights.

### **Literature Review**

Several authors have defined reliability from various angles. From Gay (1980) and Davis (1998) the study deduced that, reliability referred to the degree of consistency of a candidate's score in a test regardless of when it was taken, and who scored it.

Walkin (1990:137) brings in the element of outcome trust when he said, "reliability is the extent to which a test is dependable and consistent." These versions bring out four variables to examinations or script reliability. These are the time factor, nature of test, the candidate and marker. This study is interested in the marker variable and theoretically fuses the other three into a constant.



It was considered prudent to focus on the marker function because its output can be manipulated by training and changes in the mode of marking. Candidates mainly complain of the marking and rarely about the other three. The marker's role of interpreting the marking scheme, decoding what is on the script, assessing it and deciding on a mark provides a critical breaking point for a marking intervention.

Frith and Macintosh (1987) identified three factors affecting individual marker's essay scoring reliability. These include any interruption during the marking of a batch of scripts, which results in different standards being applied before and after the interruption break. The second factor is a change from a very bad to a good or script and lastly the subconscious influence of marker knowledge of the candidate, experience and subject content.

One can be comforted by the fact that, at university level candidates use their registration numbers instead of names, which reduces the effect of knowing the candidate. It leaves interruptions from cell phone rings, conversations and tea breaks affecting reliability. The effect of changes from very good or very bad scripts could be improved by the following strategies suggested by Bray and Jones (1992). First is the use of a refined marking scheme which in theory, "should enable a marker to evaluate the responses as fairly and accurately as possible" as argued by Walklin (1990:162). Desforges (1990:9) stressed the important role of the marking scheme when he said, "each marker should be issued with a marking scheme indicating how and for what to allocate marks." These can be useful in subjects where the content requires factual rather than expressive presentation like essays.

Examiners can't take solace in the marking scheme because of its dependence on the marker. Arguments for this view include these: first, marking schemes are structured by markers and open to marker interpretation. Second, examiners differ in experience, subject matter background and temperament on the particular day and time that script is being marked. Third, that some markers are more willing to award candidates the benefits of doubt than others. Lastly, but equally important, that some markers succumb to flowery language and neatness of writing on the expense of factual information, application, relevance and critique.

The other two strategies which Bray and Jones (1992) suggested to improve marker reliability include, having the script marked by two or three independent markers then award the average score and the use of

moderation. This strategy is applied more often in research projects moderation but require more time and personnel. A clear implication of this strategy is that, candidates have to provide more than one typed copy of the assignment or examination script. This would be very difficult and costly in Zimbabwe where the economic hardships are hitting hard on the university student population. The other limitation is that there are no lecturers to mark the scripts for a second or third time. It is common for a lucky university department to have one lecturer for a specific course. Actually some university departments are operating on part time lecturers.

The literature alluded to shows that, examinations reliability is desirable. In practice marker reliability cannot be done away with but reduced as pointed out by Bray and Jones (1992: 133) who concluded by saying, "however hard teachers try, they will not all mark the same-even with the best drawn-up marking scheme." They were supported by Curzon (1993:347) who declared that, ".....there can be no real uniformity in marking standards, since scorer objectivity is impossible, standards vary of necessity, from one examiner to another." The literature is very dumb and mute on anything related to question specialized marking and specifically any efforts to improve marker reliability in Zimbabwe's universities.

## **METHODOLOGY**

### **Research Design**

The study adopted a quantitative quasi-experimental design. This was influenced by the aim of the study (establishing the effect of a practical intervention), which requires statistical analysis, a control and experiment group for comparison. Natural setting constraints in which there was no choice of markers expect those available and different live scripts used eliminated essential ingredients of a true-experimental design as required by Kothari (1990).

### **Population and Sampling**

The population of this study is composed of 867 undergraduate end of semester Research Methods and Statistics examination scripts. A sample of 40 scripts ( $n = 40$ ) was drawn and distributed to the four markers using systematic sampling in which marker A, got scripts number 1, 5, 9, 13 and 17. Candidates were seated in alphabetic order which provided variable linear dependence, one can safely conclude that systematic sampling allocated an equal share of the talent to each marker hence population normality assumed. The number of scripts  $n = 40$ , is statistically large enough for the variables to be normally distributed and finding generalized to similar universities.

## **PROCEDURES**

### **Marking Guide Structuring**

1. A marking guide was structured by the five lecturers at a standardization meeting.
2. Each marker marked five scripts to assess the suitability of the guide in terms of student response coverage and mark allocation
3. Members meet again after marking five scripts, to discuss and improve the guide
4. Five selected scripts were photocopied so that markers had identical scripts. These were marked and discussed to synchronise interpretation of the marking guide.

### **Whole Script Marking (Control Group)**

1. The sampled 40 scripts were divided into two groups by use of a calculator random numbers matched with the last three digits of the candidate's registration number.
2. From the control group, each of the markers got five scripts to mark all the questions and record the time that he/she took to mark them.
3. All the twenty scripts were moderated by the subject coordinator who recorded the agreed absolute deviations for each script by question for each marker, as shown in Tables 1 to 4.

### **Question Specialized Marking (Experiment Group)**

1. Recorded deviations were analysed to determine marker strength. Each marker was assigned to mark the question in which he/she had the least sum of absolute deviations.
2. Markers sat at a round table; each marker had five scripts to start with.
3. Each marker marked only the assigned question.
4. Scripts rotated from marker to marker until all the four questions were marked. The last marker checked that all questions were marked and added the total marks.
5. Each marker recorded the time that he/she took to mark the twenty scripts.
6. Scripts were moderated and the agreed absolute deviations were recorded as shown in Table 5.

### Data Analysis

The indicator for inter-marker reliability (accuracy, thoroughness and objectivity) is the recorded absolute deviation for each question and marker. These are presented in Tables 1 to 4 and tables 6 (a) to 6 (C: 2) for comparison. Absolute deviations and times taken to mark each set of twenty answers in the control and experiment group were compared using t-test for paired observations to test the formulated hypotheses. Summarized data was categorized and presented in the form of frequencies to facilitate the use of Chi-square for testing hypothesis at 5% level of significance.

### Findings and Discussions

Findings are presented in the order of the research questions to facilitate linkages.

#### Organizing Question Specialized Marking (QSM).

1. There should be two or more lecturers for that particular course to mark the assignments or examination scripts.
2. Each lecturer can contribute a question and model answer for the assignment or examination paper.
3. Marking team members should discuss the marking scheme thoroughly.
4. Identical dummy scripts should be used and variations in student approaches and examiner perspectives discussed and synchronized.
5. Each marker should mark only one question in all the scripts.
6. Markers should mark from a conducive environment, probably with a low music background.

#### Effects of Question Specialized Marking (QSM)

The tables presented below, show the question, marker deviation and time analysis for experiment and control group.

#### Whole Script Marking (Control)

Marker A

Table 1 n = 5

Script Number	Absolute deviations				Totals
	Q1	Q2	Q3	Q4	
1	1	4	5	2	12
5	0	1	0	0	1
9	0	3	3	2	8
13	1	4	1	0	6
17	2	3	1	0	6
Totals	4	15	10	5	34

**Marker B****Table 2 n=5**

Script Number	Absolute Deviations				Totals
	Q1	Q2	Q3	Q4	
2	0	2	3	0	5
6	3	0	1	0	4
10	1	4	2	1	8
13	0	0	2	1	3
18	1	3	2	2	8
<b>Totals</b>	<b>5</b>	<b>9</b>	<b>10</b>	<b>4</b>	<b>28</b>

Time taken = 150 minutes, marker understands Q4 (Statistical analysis) and Q1 (Literature Review). Can Mark Q4

**Marker C****Table 3****n=5**

Script Number	Absolute Deviations				Totals
	Q1	Q2	Q3	Q4	
3	0	2	1	1	4
7	1	1	0	1	3
11	0	0	3	0	3
15	0	3	2	2	7
19	3	4	2	1	10
<b>Totals</b>	<b>4</b>	<b>10</b>	<b>8</b>	<b>5</b>	<b>27</b>

Time taken = 200 minutes. Marker understands Q1 (Literature Review) and Q4 (Statistical Analysis). Marker has problems with Q2 (Computer Application). Mark Q3

**Marker D****Table 4****n=5**

Script Number	Absolute Deviations				Totals
	Q1	Q2	Q3	Q4	
4	1	2	3	1	4
8	2	2	1	0	3
12	1	3	5	1	3
16	0	0	1	1	7
20	2	2	0	3	10
<b>Totals</b>	<b>6</b>	<b>6</b>	<b>8</b>	<b>6</b>	<b>28</b>

Time taken = 195 minutes. Marker understands Q1 (Literature Review) Q4 (Statistical Analysis). Marker has difficulties with Q3 (Project Methodology) to mark Q2.

From considering all markers, one can notice that, first Research Methodology and Computer Application had problems. The diversity of responses given required a good content base for the examiner. It is also of great interest to note that, markers A and D had more deviations; it was their first time to mark at university level. The deviation analysis has also worked well to detect marker-training needs.

## Question Specialized Marking (QSM)

Table 5 Experimental Group

Script Number	Absolute Deviations			
	Marker A	Marker B	Marker C	Marker D
21	0	1	2	1
22	2	0	1	2
23	1	0	3	0
24	0	1	0	0
25	0	2	1	0
26	0	2	1	1
27	2	1	2	0
28	3	3	0	2
29	0	0	0	5
30	1	0	0	2
31	0	0	1	1
32	0	0	0	0
33	1	3	0	7
34	0	0	1	3
35	2	0	2	3
36	3	1	3	0
37	1	0	1	0
38	0	0	1	1
39	2	1	0	2
40	0	2	3	1
N = 20	$\Sigma d = 18$	$\Sigma d = 17$	$\Sigma d = 22$	$\Sigma d = 31$
Time (Minutes)	T = 210	T = 130	T = 190	T = 160

If one holds marker variable constant, and compare scripts one can deduce that candidate 32 is good in answer presentation. His/ her responses are clear, no marker deviates with such a script. On the other hand candidate 36 had deviations from 3 markers and only in the mathematics answer did he/she not have a deviation. The analysis can be useful for teaching purposes hence should be used in marking assignments.

## Comparative Summary

Table 6 (a) Total Deviations Control and Experimental group Results

Method	Marker A	Marker B	Marker C	Marker D
Whole Script	34	28	27	36
Question Specialization (QSM)	18	17	22	31
Differences (d)	16	11	5	5

From the data in the table, following summaries can be deduced:  
 $\Sigma d = 37$ ,  $\Sigma d^2 = 427$ ,  $(\Sigma d)^2 = 1369$  and that the Mean deviation = 9.25.

A t-test for paired observations was carried out to find if the observed differences between whole script marking method and Question Specialization are significant.

### Hypothesis

$H_0$ :  $d = 0$  (Question specialization has no effect on inter-marker reliability)

$H_1$ :  $d > 0$  (question specialization improved inter-marker reliability)

From table 6(a) data:  $n = 4$ ,  $v = 3df$ , at 5% level of significance and one tailed test, tabled value of  $t = 2,353$ . The calculated test statistic for  $t = 3,481$ . Since tabled  $t = 2,353 < \text{calculated } t = 3,481$ ; Reject  $H_0$  and conclude that, Question Specialized Marking improves inter-marker reliability significantly at 5% level. In other words, the reduced deviations indicate an improvement in inter-marker reliability.

Spearman's rank order correlation co-efficient was calculated to find how, while script and Question Specialized Marking were related. The analysis produced these:  $\sum d^2 = 6$  and  $R = 0,4$  showing a weak positive correlation between the two marking methods. This could be explained in terms of marked scripts homogeneity and marker consciousness.

**Table 6 (b) Time Taken Control and Experimental Group**

Marking Method	Marker A	Marker B	Marker C	Marker D
Whole Script	250	190	200	190
Question Specialized (QSM)	210	130	190	160
Differences	40	60	10	30

From this data:  $\sum d = 140$ ,  $(\sum d^2) = 19600$ ,  $\sum d^2 = 6\ 200$  and Mean Deviation = 35  
A t = test for paired observations was carried out to find if the observed time differences are significant.

### Hypothesis

$H_0$ :  $d = 0$  (Question Specialization has no effect on marking time)

$H_1$ :  $d < 0$  (Question Specialization reduces marking time).

From table 6 (b) data:

$N = 4$ ,  $V = 3df$ , at 5% level of significance for a one tailed test, tabled value of  $t = 2,353$ .

Calculated test statistic,  $t = 3,367$ . Since tabled  $t = 2,353 < \text{calculated } t = 3,367$ : Reject  $H_0$  and conclude that, Question Specialized Marking reduces marking time significantly at 5% level.

### Hypothesis

$H_0$ :  $d = 0$  (Question Specialization has no effect on marking time)

$H_1$ :  $d < 0$  (Question Specialization reduces marking time).

**Table 6 (C: 1) Observed Deviation Frequencies**

Marking Method	Marker A	Marker B	Marker C	Marker D	Totals
Whole Script	34 (30.5)	28 (26.4)	27 (28.8)	36 (39.3)	125
Question Specialized (QSM)	18 (21.5)	17 (18.6)	22 (20.2)	31 (27.7)	88
Differences	52	45	49	67	213

To find out if marking method depends on a particular marker, a Chi-square test was carried out at 5% level of significance. Expected deviation frequencies are in brackets in table.

### Hypothesis

$H_0$ : Marking method is independent of marker

$H_1$ : Marking method depends on particular markers

From tables 6 (C: 1) and 6 (C: 2), at 5% level of significance,  $V = 3df$ , the tabled Chi-square value  $\chi^2 = 7.815$ . Calculated Chi-square 7.815. Calculated Chi-square  $\chi^2 = 2.15$

Since calculated  $\chi^2 = 2.15 < \text{tabled } \chi^2 = 7.815$ , Accept  $H_0$  and conclude that, marking method can be used by any other marker.

### Participants' Views on Question Specialized Marking

After the marking session, five participating lecturers identified the following benefits of Question Specialized Marking (QSM).

- There is maximum and efficient use of marker expertise. Each team member's special task-related skills are acknowledged and tapped for use. In this case:  
Marker A – Literature Review, Marker B – Statistical Analysis,  
Marker C – Research Project Methodology and M a r k e r  
D – Computer Application
- Question specialization mode of marking reduces marking time significantly, is through and objective, hence improves inter-marker reliability.
- Marking team is responsible for whole task – marking all examination results.



- The method has inherent internal self-supervision mechanisms. Each member works so that he/she does not delay others in the marking process.
- Each team has the autonomy to determine work method scheduling when to meet, where and how many scripts to mark per day or session.
- Question specialization method provides script security as all marking is done under one roof and mark addition checks are done by a second person.
- This method satisfies markers social needs of belonging – “people are social animals; they have a basic need to interact with others” as stated by Greenberg and Baron (1998:252).
- Question specialized marking has a high probability of reducing examiner stress as a result of marker interactions.
- This method reduces marker competition, to mark more than others and promotes co-operation, marking like others.

### **Weaknesses**

The few constraints attributed to question specialized marking have more to do with the organization than the actual marking.

- Markers are not ready to consult each other as they do when using whole script marking method.
- Lack of sufficient team member training affects marker assessment of candidates' diversified answers.
- Administration of rewards and punishments (praises, payments) linked to individuals and not the team, affects member cooperation.
- Some members' unwillingness to cooperate with each as a result of unsettled personal squabbles affects team spirit.
- Team's failure to receive support from management in the form of refreshments and background music.

### **CONCLUSION**

From the above findings, the study concluded that, the success of examinations Question Specialized Marking (QSM), depends largely on good team member management. Team members specialization should be identified from a relatively large sample of marked scripts. The QSM marking mode can be more effective if markers work in a relaxed

environment with a low music background. Question Specialized Marking reduces marking deviations and marking time significantly. It is applicable when two or more examination questions are answered and two or more markers are available for the marking task.

## **RECOMMENDATIONS**

There are more advantages than disadvantages in the application of examination Question Specialized Marking (QSM). Statistically, the method was found to improve inter-marker reliability significantly at 5% level. The study recommends that all universities with large enrolments and more markers try to apply (QSM) during the next examination sessions. ZIMSEC can also adopt the method. Further research is required to establish Question Specialized Marking's effectiveness with larger groups and different subjects.

## REFERENCES

- Bray, E. and Jones, R. L. (1992) *Assessment: From Principles to Action*, Hong Kong: Routledge.
- Bukenya, M. (2006) Comparing the reliability of the Conveyor Belt Marking system with the Traditional marking system. [0] <http://www.iaea.info/documents/paper>. accessed 29/12/2011
- Chombo, I. (2000), "Higher Education and Technology in Zimbabwe" *Zimbabwe Journal of educational Research*. 12(3). 7- 26
- Curzon, L.B. (1993) *Teaching in Further Education*, London: Cassell.
- Desforges, C. (1990). *Testing and Assessment*, London: Cassell
- Frith, D.S. and Macintosh, H.G. (1987). *A Teacher's Guide for Assessment*, London: Stanley Thorne.
- Greenberg, J. and Baron, R.A. (1998). *Behaviour In Organisations*, Toronto: Prentice-Hall.
- Kothari, C.R. (1990). *Research Methodology*, New Delhi: Wishwa.
- Malmberg, T.D. Wanner, B. Sumra, S. and Little, T.D. (2000), "Agency beliefs about School Achievement" *Zimbabwe Journal of Educational Research*. 12 (2), 126–150.
- Moberly, W. (1998). *The Crisis In The University*, London: Christian Fronter Council.
- Mutumbuka, D. (1981) "Education – Past, Present and Future: Keynote Seminar Speech," *TEACHERS FORUM*, 9 (11), 4–10.
- Okelowange, P.J. (2004a). Report of the pilot study on the implementation of conveyor Belt system of marking in Uganda.
- Walkin, L. (1990). *Teaching and Learning In Further and Adult Education*, London: Stanley Thornes.
- Zindi, F. (1998) "Arguments for and Against University Expansion" *Zimbabwe Journal of Educational Research*. 10 (1), 33–49.



This work is licensed under a  
Creative Commons  
Attribution – NonCommercial - NoDerivs 3.0 License.

To view a copy of the license please see:  
<http://creativecommons.org/licenses/by-nc-nd/3.0/>

This is a download from the BLDS Digital Library on OpenDocs  
<http://opendocs.ids.ac.uk/opendocs/>